LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-430828

The Cyber Science and Security Institute

J. M. Brase, C. W. Spain

May 11, 2010

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

V 1.0

The Cyber Science and Security Institute

Complex Information System Simulation and Analytics at Scale

James M. Brase, C. Wes Spain
Lawrence Livermore National Laboratory

What we propose

We propose the establishment of a technical center applying some of the nation's premier high-performance computing and analytic capabilities to solutions for critical national security problems in large-scale cybersecurity. The center will bring together a partnership of DOE National Laboratories, industry leaders, and research universities to address cybersecurity R&D topics including:

- Modeling and simulation of complex networks at Internet scales to enable high-fidelity test, evaluation, and training
- Developing real-time network situational awareness through pattern discovery in massive data streams
- Automated analysis of full-scale software systems to discover subtle vulnerabilities

The institute's computing and research facilities and R&D programs will be open collaborations where researchers from the Labs work side-by-side with industry and academic experts to develop and test new approaches to building both near-term solutions and long-term scientific foundations. The proximity of the institute to classified National Laboratory programs will provide clear transition paths into national security applications. The institute will operate by developing and evaluating prototype systems and methods at full-scale and working with government sponsors and industry partners to apply the resulting solutions to real-world operations.

Why is supercomputing important for cybersecurity?

The world relies on complex interconnected information networks operating at incredible scales and speeds that are continually increasing. The ability to interact with confidence and well-understood trust relationships on these networks is critical to both business operations and to national security. The nation has an opportunity to build the science and security foundations of these networks that effectively leverages the major investments in high-performance computing at the National Labs, the expertise established in university programs for building network simulations and testbeds, and deep industry expertise and resources in computing and networking. The computing approaches discussed here do not address every problem in cybersecurity but are an important and currently underutilized

tool. Sustainable security "at scale" with measurable improvements will require in-depth understanding of large-scale networks and new high-performance computing tools beyond those available today.

To operate securely network operators must be able answer basic questions:

- What is currently happening on our network?
- If the system changes in a particular way, are more secure or less secure?

Today, we cannot answer these questions at the scales, speeds, and complexities of real networks. High-performance computing is one of the critical tools for developing the approaches that can ultimately address these questions.

The proposed Institute will focus on three areas of cybersecurity where supercomputing can enable major gains in capability and directly impact national security:

- **Complex network modeling and simulation** - Test and evaluation for cybersecurity at large scales is very limited today. It is impossible to commandeer large networks for testing "in the wild" and dedicated and isolated test ranges are limited to a few hundred nodes. Developing simulation capabilities that allow scaling to very large networks while maintaining required fidelity is critical to sustainable progress in cybersecurity. These simulations will stress the largest computing capabilities. Simulating a day of operation of a million node network requires simulating billions of messages between computers and will require petaflop class machines for solution in useful times.

- **Network situational awareness** - Requires building sophisticated statistical models based on millions of communication events per second on links connecting millions of network nodes. Today these analyses are performed by human analysts usually weeks after the data is recorded. Real-time discovery of patterns in these large-scale data streams will require innovative algorithms operating on high-performance computer architectures.

- **Scalable software vulnerability analysis** - Deep analysis of all possible vulnerabilities in million-line software systems requires evaluation of trillions of possible execution paths in the codes. Current tools are limited to small software systems or very shallow analyses. Extending these capabilities to the largest supercomputing systems will create unique capabilities in analysis of large-scale software for government and critical infrastructure systems.

The difficulty in advancing cybersecurity is not due to the lack of good ideas. The problem is the inability to evaluate these ideas quantitatively, eliminate the bad ones, and optimize the good ones under realistic conditions. Cybersecurity R&D at the scales, speeds, and complexities needed for today's network defense is a national need that is currently not adequately addressed. Bringing some of the nation's premier supercomputing capabilities and tools to bear on this problem is the objective of this proposal.

The Institute and its functions

The Institute will be based at a new open campus adjacent to LLNL and Sandia with easy access for industry and academic partners at an industry-standard level of security. It will include high-performance computing and data analysis laboratories in a collaborative R&D space for Lab and partner researchers. The Institute will be organized at three levels of functionality.

Physical level – Partnering to create a scalable cyber lab
The partners will work together to create a set of computing and analysis resources at scales beyond what could accomplish individually. This model was successfully developed and tested in the Hyperion supercomputing partnership. We propose to adopt that approach for building the computational resources for the Institute. Dedicated laboratory facilities (e.g. a building) in the Livermore Valley Open Campus will require a funding sponsor. Each partner will have full access to the resources at this level in proportion to their contribution to the resource.

Computing and cyber research facilities and expertise inside LLNL and Sandia will be accessible to center programs that require classified activities or connection to existing Laboratory programs. Facilities at partner institutions can be remotely linked into the center to enhance partnership interactions.

Simulation and analysis framework level – Collaborative tools supporting R&D
The Institute will develop an open framework for large-scale network simulation and analysis utilizing the computational resources of the physical level. This framework will extend on-going work at the Laboratories and partner companies. The intent is that this framework will be openly available to all the partners. The framework will support large-scale packet-process level simulation integrated with extensive virtualization capability. To be successful, any simulation program must be accompanied by a parallel network measurement program that provides models of networks to be simulated and data for validation of the models. This parallel structure keeps the simulation efforts grounded in real networks and builds in validation from the start.

Focused R&D project level – Developing new cybersecurity capabilities
At the top level are a set of specific R&D projects that make use of the simulation and analysis framework and the underlying computational resources. Areas of particular interest include real-time network situational awareness, analysis of the security properties of large software systems, and test and evaluation of network defense tools and policies. These projects will have multiple funding sources and can involve individual partners or subsets in collaboration.

Building the future workforce needed for secure networks is an important national need and a central function of the center. Working with our partners the center will establish educational programs providing opportunities for research in new cybersecurity activities and training with hands-on experience with state-of-the-art facilities and experienced practitioners.

Building the Partnership

The partnership proposed here brings together extensive experience in high-performance computing and simulation, large-scale networking, and in developing and operating cyber security test beds and development environments. Discussions have included

- DOE National Labs – LLNL, Sandia, ORNL, PNNL
- Industry – Cisco, IBM, Intel, Red Hat, Dell-Perot
- Universities – USC/ISI, UC Berkeley, UC Davis, Carnegie-Mellon

The institute will build upon LLNL's Hyperion partnership approach in which each partner contributes to the common facilities and programs while receiving a share of access to the laboratory resources. This approach takes advantage of the large investments made in high-performance computing and cyber security at the National Labs and in university programs while creating a strong tie to private sector capabilities and resources. We envision the partners having a physical presence with people located at the institute while also participating in institute activities through virtual connections.

The partners in the center have roles in at least three areas:

- First, the partners are central to standing up the capabilities of the center. In return they benefit from access to the larger set of shared resources at both the physical computing level and at the simulation and analysis framework levels;
- Second, the partners are integral components of the cybersecurity and computing R&D programs executed in the center;
- Third, the partners provide a natural technology transfer path for the R&D products of the center into government and private sector applications.

Finally, the government sponsors of the work in the center are an important component of the partnership. Active sponsor participation, including resident people, will increase the effectiveness of center programs.